

SAMGrid Road Map

Adam Lyon
GDM - 2006 Feb 28

February activities highlights (see status doc)

- ◆ v7 in use on CDF Farm
- ◆ DØ Refixing complete
 - ❖ All the I/O of the yearlong p17 reprocessing project squeezed into 6 weeks!
 - ❖ ~50% offsite with SAMGrid
 - ❖ ~10% with LCG+OSG resources
 - ❖ A triumph for SAM (job + data handling)
- ◆ Fixed elusive and annoying DB Server problem at CDF
 - ❖ Running out of file descriptors
 - ❖ Traced to OmniOrb bug (fixed in latest version)
- ◆ 10 submissions to CHEP 2006! Many on lessons learned.

February activities lowlights (see status doc)

- ◆ FSS problems made CDF roll their own storage system
- ◆ Stupid “misunderstandings” in v7 DØ sam-manager
 - ❖ Incorrect file type assignment
- ◆ DB Server memory explosion seen at DØ
 - ❖ Cannot reproduce
 - ❖ Not seen at CDF
 - ❖ ???

Road Map

- ◆ Available human resources
- ◆ Where is the project going in the short and longer term?
- ◆ How do we prioritize?
- ◆ What do we do if the landscape changes?

The people power

- ◆ 100%: Andrew, Parag, Steve Sherwood
- ◆ 50%: Randolph, Steve White, Robert Illingworth, Dehong, Krzysztof, myself
- ◆ 20% Gabriele

- ◆ ~ 6 FTE's

Effort	FTE
Core Development	2.5
Deployment to Production	1.3
Operational Support	1.0
Project Management	0.5
Outreach	0.5
Total	5.8

Continue smooth operations

- ◆ Expert support of SAM DH and SAMGrid
- ◆ Top priority task – if we fail here, the project fails
- ◆ But can be major disruptions – unplanned
- ◆ Why does SAM still require expert support (why do we still find bugs)?
 - ❖ While our testing is improving, we cannot reproduce the production environment
 - ❖ Introduction of multithreading adds complications we are still learning how to handle
 - ❖ Limited ad hoc monitoring
 - ❖ Installation/configuration were designed to be flexible, not easy
 - ❖ CDF and DØ have different load levels and usage patterns. They exercise the code differently. They hit different problems.

... continue smooth operations

- ◆ Anecdotal evidence that our steady state operations load is decreasing
 - ❖ SAM still functions, even with the loss of major players (Sinisa, Lauri, Valeria – to their credit)
 - ❖ While the support load is large, we are still able to get SAM tasks completed
- ◆ Everyone works on operations
 - ❖ SAM Station+FSS/C++ API: Andrew
 - ❖ SAMGrid: Andrew, Parag, future DØ “camper”
 - ❖ DB server: Steve W, Randolph
 - ❖ Python client: Robert, Steve S.
 - ❖ DØ: Robert, future Dehong; CDF: Dehong, Randolph

Near term tasks

- ◆ Upgrade to Python 2.4
 - ❖ Client already there
 - ❖ Problems with DB Server

- ◆ DØ Upgrade to v7
 - ❖ SAMGrid, Online, MC Generation, Users

- ◆ Complete deployment at CDF
 - ❖ Automated job restart, “sam get dataset”

- ◆ MIS
 - ❖ New monitoring system long time in the making
 - ❖ Now testing at the multi-server level
 - ❖ DB retention policy
 - ❖ SAM HDTV is already working

... Near term tasks

- ◆ SQLBuilder
 - ❖ Replacement for unmaintainable dimensions parser
 - ❖ Needed by experiments for enhanced queries
- ◆ Improve testing capabilities and documentation
 - ❖ We have good tests of the DB server
 - ❖ But we need specific client tests,
 - ❖ Testing of autodeestination
 - ❖ SAM station tests
- ◆ Testing for Oracle 10g

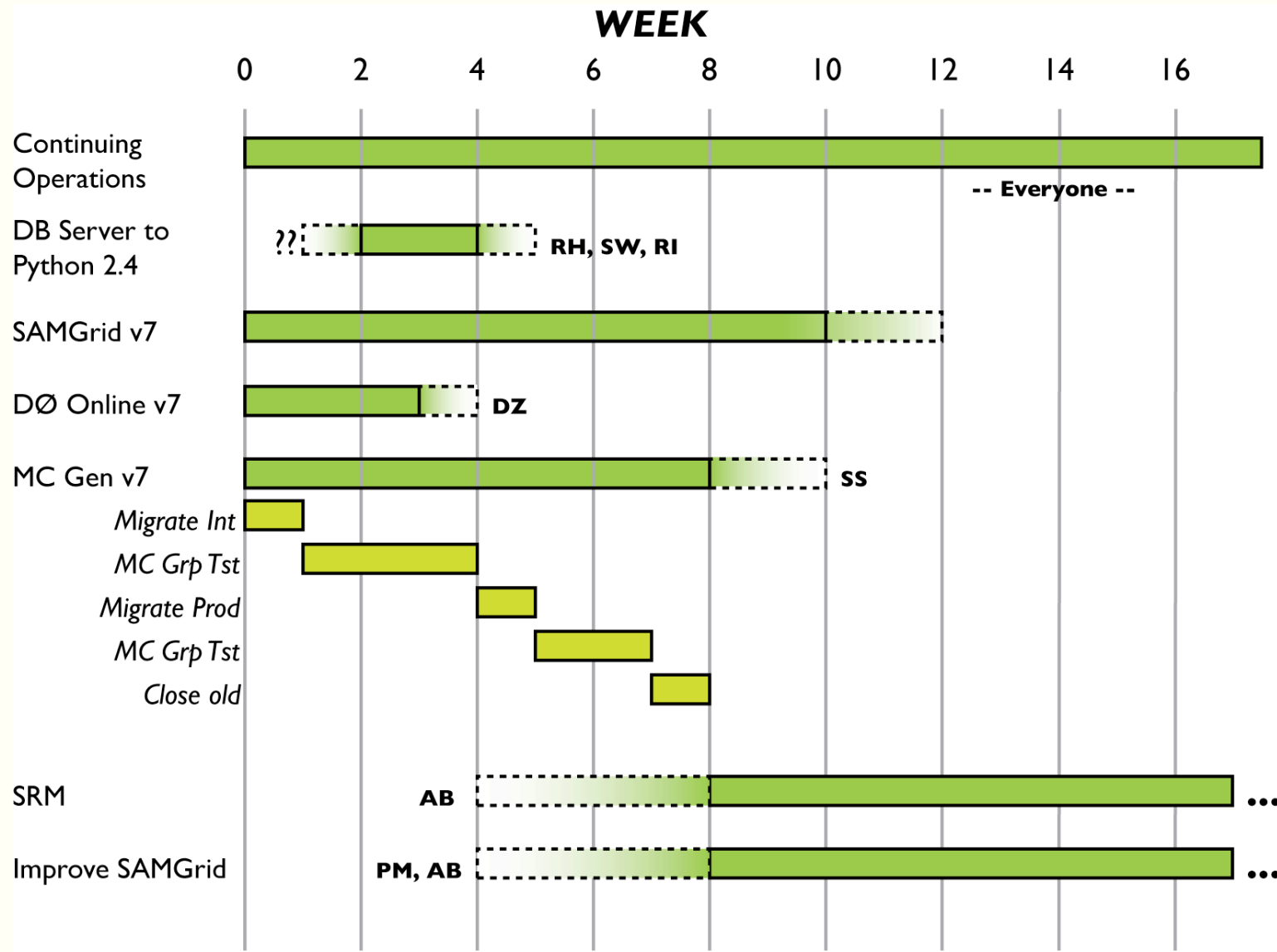
Longer term

- ◆ Improved monitoring (cache metrics)
 - ❖ Make use of MIS
- ◆ Improved SAMGrid performance, deployment, stability
- ◆ SRM interface
 - ❖ Essential for access to dCache and for running on the Grid (LCG, OSG, glide ins)

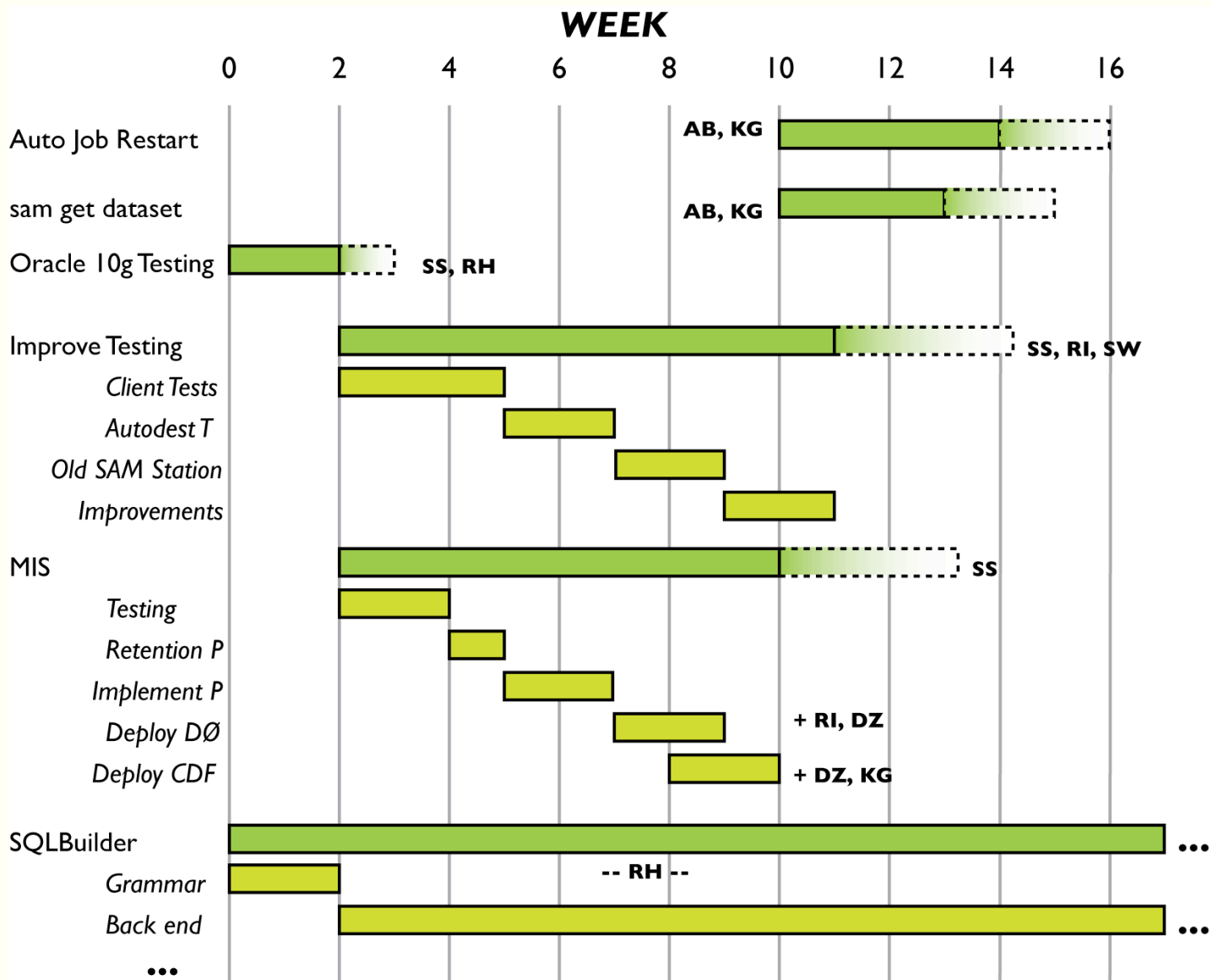
Longest term

- ◆ SAMGrid for analysis jobs
- ◆ Breakup of SAM into individual service

Timeline



... timeline



Priorities of near term tasks

1. Operations support [if we do not support our products, we fail]
2. Upgrade to Python 2.4 & Oracle 10g [known problems with Python 2.1, the upgrade to Oracle 10g is mandatory]
3. DØ v7 upgrade; Improved testing/docs
[Without these, SAM can still function, but experiments will suffer, we will lose already invested work, and our operations will not decrease]
4. Automated job restart; “sam get dataset”; MIS, SQLBuilder
[SAM will continue to function without these, but at perhaps a compromised level and not meeting experiments requirements; lose already invested time and work]

Future priorities

1. Improved monitoring, SAMGrid performance/deployment/stability
[SAMGrid can function without these tasks, but at a higher operations level]
2. SRM Interface
[SAM works now without SRM interface, but as the Grid becomes more prevalent, experiments will need to find an alternative to SAM to make use of storage elements; CDF will remain with the ad hoc dCache station]
3. SAMGrid for analysis
[DØ will need to find an alternate to SAMGrid for running user jobs on the Grid]
4. Break up SAM into services
[SAMGrid development stops]

Risks and contingencies

- ◆ Unplanned tasks appearing
 - ❖ Refer to GDM for evaluation and approval
- ◆ If a task gets into trouble, a persons from a lower priority task could help (but reality is that people are too pigeon holed)
- ◆ If a drastic cut needs to be made, the most vulnerable near term tasks are MIS and SQLBuilder. Could forgo some testing, but operations would not decrease
- ◆ The future of SAMGrid is also vulnerable.
 - ❖ SRM is essential for DØ and CDF to fully utilize the Grid
 - ❖ SAMGrid for analysis may be up for debate
 - ❖ Breaking up SAM depends on how far we want to take the project and the position the CD desires to have in the Grid world.